2007 ITE Annual Meeting
Pittsburgh, PA
August 6, 2007

Session 14: Strategies for Signal System Control
Track: Traffic Engineering/Multimodal Strategies

"Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion"

Jim Clark, P.E.
Tindale-Oliver & Associates, Inc.

"Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion"

Purpose:

Focus on the objective function for traffic signal timing split optimization and present the comparative results of a related simple signal timing experiment.

Note: This is **not** about comparing specific split optimization software methods being applied today.

"Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion"

Some methods used to develop optimum signal timing splits

- A traffic cop
- "Seat-of the-pants"
- Field trial and error
- Formula calculation by hand….mathematical deterministic expressions.
- Computer software………….mathematical and subjective expressions.

"Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion"

"What is the basis of the split optimization function being applied?"

…………most traffic engineers simply do not know.

Yet, without:

- *understanding the basis of objection function and*
- *traffic performance measure being used to determine timings*

how do we know we are producing sensible optimum timing splits?

"Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion"

Current methods can be broadly categorized in either one of two forms, based on optimization function's focus:

- **Volume to Capacity (V/C) Methods**

  Based on some form of V/C ratio, usually to either:
  - balance the V/C of critical movements, or
  - theoretically approximate minimization of total intersection delay.

- **Movement Delay Methods**

  Based on specific evaluation of:
  - individual movement's delays and
  - usually focused on minimizing critical movements' delay

"Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion"

These two methods are **profoundly different** approaches to determining optimum signal splits and later a simple **experiment** will clearly **illustrate differences**.

Notes:

- Methods that explicitly optimize total intersection delay (or something similar), for this paper, fall into V/C Methods category.
- Methods that might not be so obvious that fall into the V/C Methods category include:
  - Critical movement Analysis (CMA) of Circular 212
  - Intersection Capacity Utilization (ICU) method
- When V/C ratios approach 1.0 or higher, any optimization objective function will have difficulty determining best splits; results observed in street will be bad regardless.

"Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion"

Volume to Capacity (V/C) Methods

- "Traffic Signal Settings", F.V. Webster, B.Sc., Ph.D., Road Research Technical Paper No. 39, London Her Majesty's Stationery Office, 1958.
- A pioneer in traffic engineering

Webster's Method

DEPARTMENT OF
SCIENTIFIC AND INDUSTRIAL RESEARCH
ROAD RESEARCH LABORATORY

Road Research Technical Paper No. 39

Traffic Signal Settings

BY

F. V. WEBSTER, B.Sc., Ph.D.

LONDON
HER MAJESTY'S STATIONERY OFFICE
1958

Webster's Method (Cont.)

SUMMARY OF PROCEDURE FOR SETTING TRAFFIC SIGNALS

In a period where the traffic flow is varying randomly about the mean, the procedure for obtaining optimum settings is as follows:

(i) Estimate the flow and saturation flow for each arm of the intersection.

(ii) Evaluate the ratio of flow to saturation flow for each arm, and select the y value for each phase (i.e. the maximum q/s value).

(iii) Add the y values together to give Y for the whole intersection.

(iv) Decide on all-red periods for pedestrians, turning traffic, etc. and estimate the lost time, R, due to this, e.g. if sequent ambers occur twice per cycle then $R=6$ seconds; if there are two all-red periods of 2 seconds each then $R=10$ seconds (see Fig. 6).

(v) Calculate the cycle time from equation (4):

$$c_0 = \frac{1{\cdot}5L + 5}{1 - Y}$$

where L is the total lost time per cycle, given by

$$L = nl + R$$

where n is the number of phases and l is the average lost time per phase due to starting delays.

(vi) Subtract the total lost time, L, from the cycle time giving the available green time and divide this in the ratio of the y values, i.e.

$$g_1 = \frac{y_1}{Y}(c_0 - L)$$

$$g_2 = \frac{y_2}{Y}(c_0 - L) \text{ etc.}$$

(vii) Add l seconds to each effective green time, $g_1, g_2 \ldots$ and subtract the amber period (3 seconds) to give the controller setting of green time.

Webster's Method (Cont.)

Equal V/C for all phases

MISCELLANEOUS RESULTS

Degree of saturation

It is shown in Appendix 4 that for optimum division of the cycle time the degree of saturation should be the same for all phases of the intersection. In this calculation, we have considered only one arm from each phase—the one with the highest q/s value. The degree of saturation for optimum settings of the controller appears to be independent of the amount of lost time per cycle, depending only on Y. It is given by equation (4·7) in Appendix 4 as

$$x_0 = \frac{2Y}{1+Y} \qquad \text{.......(6)}$$

Average delay for the whole intersection

The average delay to all vehicles using an intersection has been deduced for optimum settings of the controller. The steps of the calculation are shown in Appendix 4 where the average delay per vehicle is given by equation (4·12) as

$$\bar{d} = \frac{c_0}{2}\left\{1 - \frac{\sum_{1}^{n'} y_r \, q_r}{YQ} + \frac{2n' Y^2}{LQ(1+Y)}\right\} \quad \text{approximately....(7)}$$

where n' is the number of approaches to the intersection. This expression applies only to junctions where all arms of any one phase have approximately the same ratio of flow to saturation flow. The expression does not include the empirical correction term of equation (1), but this can be taken into account, approximately, by reducing $\bar{d}$ by about 10 per cent.

Webster's Method (Cont.)

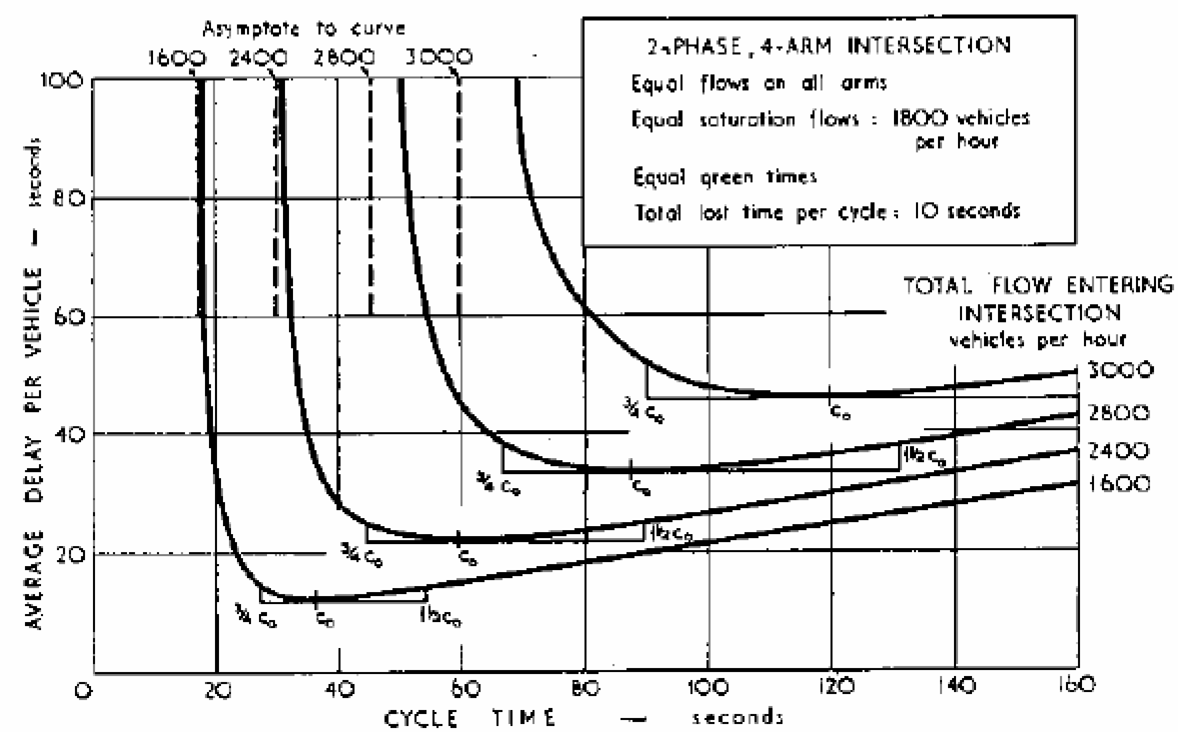


FIG. 7. EFFECT ON DELAY OF VARIATION OF THE CYCLE LENGTH

From graphs such as Fig. 7, it was found that in most practical cases the delay for cycle times within the range $\frac{3}{4}$ to $1\frac{1}{2}$ times the optimum value is never more than 10 to 20 per cent greater than that given by the optimum cycle. This fact can be used in deducing a compromise cycle time when the level of flow varies considerably throughout the day. It would be better either to change the cycle time to take account of this, or, as is more common, to use vehicle-actuated signals. However, for a single setting of fixed-time signals the simple approximate method outlined below may be used.

Summary of Webster's Research of Volume to Capacity (V/C) Methodology

Most important observations of Webster's research

- Explicit objective – is to **equalize V/C ratio** for the critical movements of all phases.
- Theoretically achieves the **least delay for "all" vehicles** using the intersection.
- **Performance measure** is critical movement V/C as a surrogate for **total intersection delay**.

"Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion"

Movement Delay Methods

- **Roots** within **Highway Capacity Manual** (HCM) procedures.
- In United States, **HCM** in one form or another is basis for virtually **all roadway capacity and operational analysis.**
- **HCM** is a long time running; **first publication some 57 years ago!**

"Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion"

Highway Capacity Manual – Long time running:

Edition	Published
1st	1950
2nd	1965
Interim	1980
3rd	1985
4th	1994
5th	1997
6th	2000

Highway Capacity Manual (HCM) – History:

Important findings:

- The 1985 HCM delay equation - was originally based on the delay concepts formulated by Webster in 1958.
- As HCM evolved, the **HCM three part delay equation,** for a given lane group, has become perhaps the **most significant tool available for determining optimum signal timing splits** (as described later).

HCM Three Part "Delay" Equation for given "lane group"

HCM 2000 Delay Equation

The average control delay per vehicle for a given lane group in the HCM 2000 is calculated by using the following equation

$$d = d_1 \times PF + d_2 + d_3 \quad (3)$$

where d = control delay per vehicle, s/veh;
d_1 = uniform control delay assuming uniform arrivals, s/veh;
PF = uniform delay progression adjustment factor, which accounts for effects of signal progression (in this paper, $PF = 1$ because an isolated intersection is assumed);
d_2 = incremental delay to account for effect of random arrivals and oversaturation queues, adjusted for duration of analysis period and type of signal control; this delay component assumes no initial queue for a lane group at the start of analysis period, s/veh; and
d_3 = initial queue delay, which accounts for delay to all vehicles in analysis period due to an initial queue at the start of analysis period, s/veh. A zero initial queue is assumed in this paper.

$$d_1 = \frac{0.50c\left(1-\frac{g}{c}\right)^2}{1-\left[\min(1,x)\frac{g}{c}\right]}$$

$$d_2 = 900T\left[(x-1)+\sqrt{(x-1)^2+\frac{8kIx}{CT}}\right]$$

where T = duration of analysis period, hour;
k = incremental delay factor that is dependent on actuated controller settings;
I = upstream filtering/ metering adjustment factor;
C = lane group capacity, vph; and
x = lane group v/c ratio or degree of saturation.

ESTIMATION OF d_3

A generalized form of d_3 appears as Equation F16-1, which provides estimation of the initial queue delay per vehicle (in seconds) when an initial queue of size Q_b is present at the start of the analysis period T. d_3 is a term in the delay model given in Equation 16-9.

$$d_3 = \frac{1800Q_b(1+u)t}{cT} \quad \text{(F16-1)}$$

where
Q_b = initial queue at the start of period T (veh),
c = adjusted lane group capacity (veh/h),
T = duration of analysis period (h),
t = duration of unmet demand in T (h), and
u = delay parameter.

The parameters t and u are determined according to the prevailing case. Equations F16-2 and F16-3 may be used to estimate the values for Cases III, IV, and V:

Relationship of Highway Capacity Manual (HCM) / Movement Delay Methods

HCM

- Historical HCM evaluated individual movement's levels of service (LOS).
- **Current HCM** evaluate **individual movement delays** to determine their LOS.

Movement Delay Methods

- Similar focus to current HCM - optimize based on individual movement delay.
- Use iterative procedures rather than formula-based calculations to determine signal splits either by:
  - Balance critical **movement delays** or
  - Prioritize certain **movement delays** (as long as other movements do not exceed allowed acceptable threshold delays).

Summary of HCM Movement Delay Methods

Important observations:

- HCM has **always** been focused on **individual movement not intersection performance**…..movement delay split optimization methods apply the same focus.
- HCM is now focused on a **measure that matters to the driver – individual movement delay, not V/C**….movement delay split optimization methods focus on individual movement delays.
- HCM has **never** focused **solely on total intersection delay** as the performance measure….. nor do movement delay split optimization methods.

"Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion"

Comparison of Methodologies:

Profound differences in using either method to determine optimum signal timings.

- **Webster's and other V/C Methods focus on:**
  - **Volume to capacity (V/C) derivative** and on
  - **total intersection delay** as the implied performance measure.
- **Movement Delay Methods focus on:**
  - (critical) **movement delay** and **not** on **V/C or total intersection delay**.

"Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion"

To Webster's Credit

- A true pioneer.
- First to introduce a system **approach** which was **rationale and feasible**.
- Invented the concept of "**delay**" that could be calculated.
- His mathematical formula that forms the basis of signal timing optimization could be solved by **using a slide rule** (no computer).
- This was fine 50 years ago, but let's look at this methodology as it applies today.

Primary Limitations to a V/C Based Methodology

- **V/C** as the 'explicit optimization variable' is the **wrong measure** of effectiveness:
  - V/C is an **irrelevant objective**.
  - A measure drivers **cannot and do not perceive.**
  - Drivers do not understand V/C ratio, their **performance measure is delay.**
  - Balanced V/C ratios:
    - o **do not mean balanced delay** or queues.
    - o **do not imply equitable assignment of green**.

Primary Limitations to a V/C Based Methodology

- **Total Intersection Delay** as the 'implicit optimization variable' masks the problem:
  - **Delay of critical movements is inversely proportional to their volumes.**
  - Low volume critical movement **delay might be 10+ higher** than the high volume movement delay!
  - High **delay** for movements with **low volumes almost disappears** from intersection total, due to small weight low volume has on total.
  - Drivers do not perceive total intersection delay, **only their own delay.**

Low volume critical movements experience unreasonably high delays

Doubtful that drivers wonder what the V/C ratio and intersection delay is?

- Drivers measure is **not V/C or intersection delay**, simply the amount of **their movement delay**.
- **Key problem** with V/C approach:
  - Equitable balance is achieved for a **variable that doesn't matter** to a driver (V/C)
  - This can result in unreasonably high values of delay to individual movements… ……….this being the **variable that does matter**.
- This is why **movement delay** is used as the **primary measure in all versions of HCM since 1985.**

“Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion”

Challenge to the profession:

- What are the split optimization objective functions?
- What should these objectives be?
- How are we going to achieve these objectives?

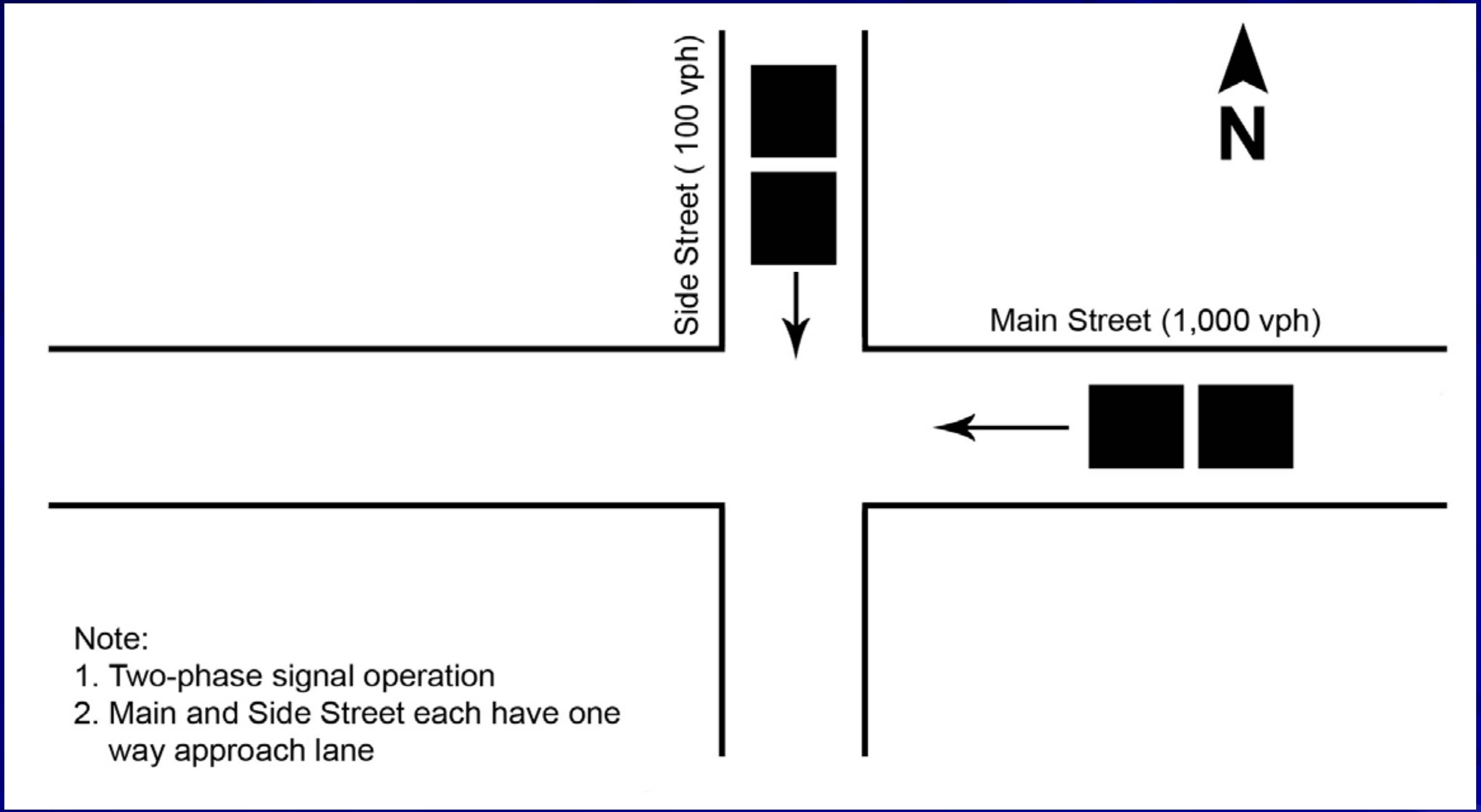
*…… again, this is **not** about software methods.*

"Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion"

Simple controlled experiment

- Results will clearly demonstrate that signal timing split optimization is best achieved and most palatable to drivers when based upon critical movement delay rather than V/C approach.
- Controlling just a few parameters; use of two dimensional tables / graphs.
- Engineer easily recognize value of critical movement delay approach over traditional V/C methodology.
- A more complicated experiment would produce quite similar results, but more difficult to see with clarity.

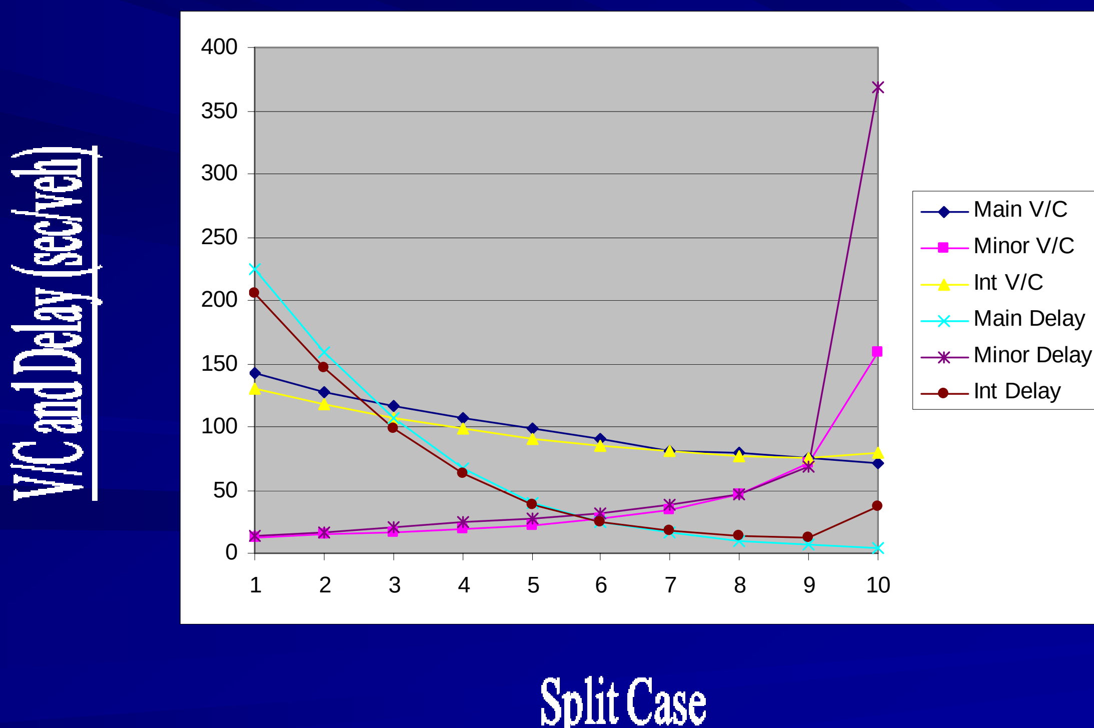
Simple Controlled Experiment



* [Cycle Length = 100 seconds, two phase operation, 3 seconds clearance time per phase; main street 1,000 vehicles per hour, one lane, one-way; minor street 100 vehicles per hour, one lane, one-way.]

HCM Critical Movement Delay Experiment

Split Case	Main Green	Minor Green	Main V/C	Minor V/C	Int V/C	Main Delay (sec/veh)	Minor Delay (sec/veh)	Int Delay (sec/veh)
1	45	49	1.42	0.13	1.30	224	14	205
2	50	44	1.28	0.15	1.18	159	17	146
3	55	39	1.16	0.16	1.07	107	20	99
4	60	34	1.07	0.19	0.99	67	24	63
5	65	29	0.98	0.22	0.91	40	28	39
6	70	24	0.91	0.27	0.85	24	32	25
7	75	19	0.81	0.34	0.81	16	38	18
8	80	14	0.80	0.46	0.77	10	46	14
9	85	9	0.75	0.71	0.75	7	68	12
10	90	4	0.71	1.59	0.79	4	369	37



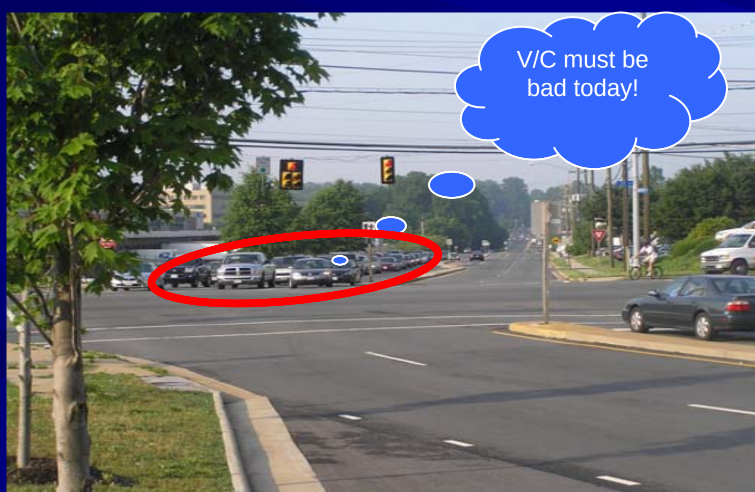
Comparative Results: V/C Methods versus HCM Movement Delay Methods

Split Case	Main Green	Minor Green	Main V/C	Minor V/C	Int V/C	Main Delay (sec/veh)	Minor Delay (sec/veh)	Int Delay (sec/veh)
1	45	49	1.42	0.13	1.30	224	14	205
2	50	44	1.28	0.15	1.18	159	17	146
3	55	39	1.16	0.16	1.07	107	20	99
4	60	34	1.07	0.19	0.99	67	24	63
5	65	29	0.98	0.22	0.91	40	28	39
6	70	24	0.91	0.27	0.85	24	32	25
7	75	19	0.81	0.34	0.81	16	38	18
8	80	14	0.80	0.46	0.77	10	46	14
9	85	9	0.75	0.71	0.75	7	68	12
10	90	4	0.71	1.59	0.79	4	369	37

Split Case	Main Green	Minor Green	Main V/C	Minor V/C	Int V/C	Main Delay (sec/veh)	Minor Delay (sec/veh)	Int Delay (sec/veh)
Equal V/C	85.5	8.5	0.75	0.75	0.75	6	74	13
Min Int delay	84.4	9.6	0.76	0.67	0.75	7	63	12
Equal Delay	67.7	26.3	0.94	0.34	0.88	30	30	30
Priority Delay	76.7	17.3	0.83	0.37	0.79	14	40	16

With V/C methods, major high volume movements enjoy disproportionately low delays….

…at the expense of the low volume movements which experience high delays….



……….even the lower volume critical turning movements exiting from the high volume roadway can experience significant delay causing traffic congestion….

…………Is this type of split optimization really sensible?

“Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion”

Conclusion

- **V/C is irrelevant and it is not an appropriate objective function** for determining optimum splits.
- **Webster-type V/C method can produce unreasonably high delays** for low volume critical movements.
- Splits based on an optimization of movement V/C or overall total intersection delay is the **wrong performance measure**.
- **HCM delay for individual movements** the is **current definitive measure of intersection performance**; split optimization should reflect this process.
- Splits should be based on individual (critical) movement delays if we are **to achieve sensible and defensible signal timings** for **all** movements.

"Assessing the Sensibility of Signal Timing Split Optimization in Addressing Congestion"

(Given the limitations of a V/C-based split optimization methodology)

The challenge to the traffic engineering professional is to determine:

- Is V/C-based methods really a valid objective for proclaiming optimum splits?
- What should the objective function be based upon – HCM movement delay?
- How is the profession going to achieve this objective?

In response to this challenge, it is recommended that ITE:

- Investigate this topic further,
- Involve appropriate traffic flow theorists, experts and practitioners to resolve this.
- Provide guidance regarding the development of optimum signal timing splits.